



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : H04M 3/50, 7/00, G06F 17/30	A1	(11) International Publication Number: WO 98/35491
		(43) International Publication Date: 13 August 1998 (13.08.98)

(21) International Application Number: **PCT/GB98/00194**

(22) International Filing Date: 22 January 1998 (22.01.98)

(30) Priority Data:

97300741.2 5 February 1997 (05.02.97) EP

(34) Countries for which the regional or international application was filed: GB et al.

(71) Applicant (for all designated States except US): BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).

(72) Inventor; and

(75) Inventor/Applicant (for US only): JOHNSTON, Robert, Denis [GB/GB]; 25 Haughgate Close, Woodbridge, Suffolk IP12 1LQ (GB).

(74) Agent: LLOYD, Barry, George, William; BT Group Legal Services, Intellectual Property Dept., Holborn Centre, 8th floor, 120 Holborn, London EC1N 2TE (GB).

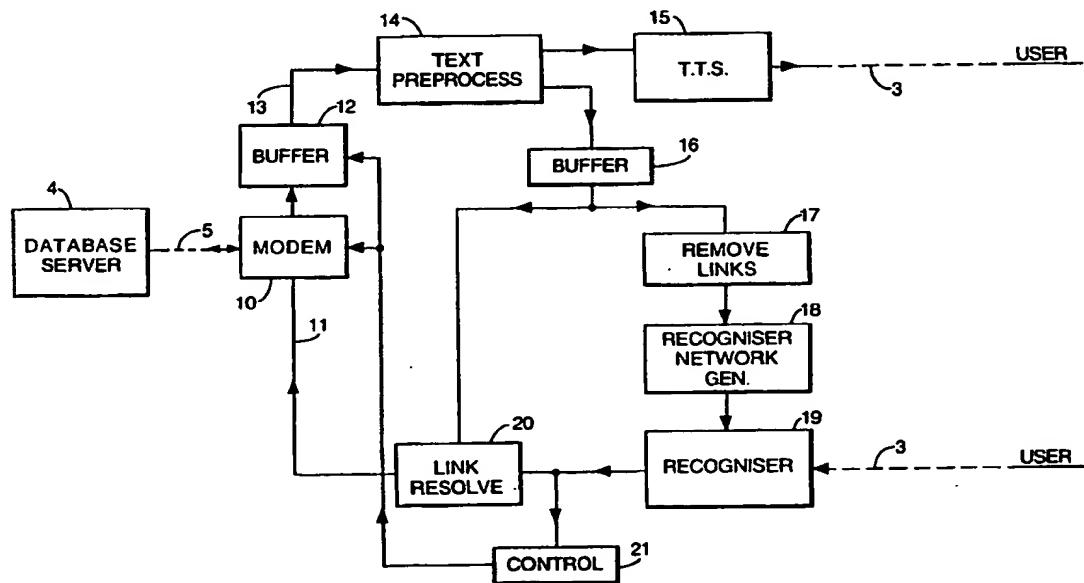
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: VOICE-DATA INTERFACE



(57) Abstract

A page of text from a database (4) has certain words marked with the addresses of other, linked pages. The text is received at (10) and converted into an audio signal by a speech synthesiser (15) that it can be heard by a user. The user's spoken responses are fed to a speech recogniser (19) so that an address associated with a marked word in which the user is interested can be returned to the database (4) for retrieval of the corresponding linked page. Because the user will not necessarily know which words are marked, the recogniser is set up to match the user's response against the whole of the text fed to the synthesiser to identify the words in the text giving the best match. A resolver (20) finds the nearest marked word to the words identified and extracts the associated link address.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

VOICE-DATA INTERFACE

The present application is concerned with voice-interactive access to text-based services.

5 According to one aspect of the invention there is provided an interface for a voice interactive service comprising:

a speech synthesiser to receive coded signals representing sequences of words and to generate audio signals corresponding thereto for output;

10 speech recognition means connected to receive the said coded signals and operable upon receipt of a speech signal to be recognised to identify that part of the word sequence represented by the coded signals which most resemble the speech signal to be recognised.

In another aspect the invention provides a method of operating a voice interactive service comprising

15 (a) receiving coded signals representing a sequence of words and synthesising audio signals corresponding thereto for output;

(b) receiving a speech signal and identifying by means of a speech recogniser that part of the word sequence represented by the coded signals which most resembles the received speech signal; and

20 (c) using the recognition result to select a further sequence of words.

Other aspects of the invention are set out in the claims.

Some embodiments of the invention will now be described, by way of example, with reference to the accompanying drawings.

In Figure 1 an apparatus 1 for providing a voice-interactive service is shown and in this example it is intended to allow a user to access a text-based information service by voice only, using a telephone 2. Although the apparatus 1 could be located at the user's premises or at the location of the text-based information service, in this example it is located at a telephone exchange or other central location where it can be accessed by many users (at different times or - 25 with duplication of its functions - simultaneously) via a telecommunications link such as a PSTN dialled connection 3. The information service is provided by a remote database server 4 which contains (or forms a gateway offering access to) 30 stored pages of textual information - though the database could if desired be

incorporated into the apparatus 1. Here we suppose that the server is part of a network accessible via a telecommunications link 5, such as the Internet, and responds to addresses transmitted to it by sending a document identified by that address. Documents provided by the Internet are commonly formatted according
5 to the hypertext markup language (HTML) which is itself a particular example of the standard generalised markup language according to international standard ISO 8879. As well as containing text characters forming the words of the text, an HTML document also contains formatting information suggesting the appearance of the document when displayed on a screen (or printed) such as position, font
10 size, italics and so forth. The precise details of these are not important for present purposes; one thing that is of significance however is that these documents also have provision for flagging words or phrases as associated with the address of another document. Part of such a document is illustrated in Figure 2a with its displayed appearance shown in Figure 2b. It is seen that this format and control
15 information is enclosed with chevrons "<>" as delimiters, not being intended for display. The text "Patent Office Sites" is to be shown in bold type as indicated by the start and finish codes and . The text "US Patent and Trademark Office" is flanked by <a> and delimiters which normally cause the text to be displayed in a distinctive manner - a special colour or underlined, for example -
20 to identify this phrase as representing a link. Moreover the <a> code contains an associated address "http://www.uspto.gov" which is the address of the Internet page of the US Patent and Trademark Office. When a user with a visual display terminal receives such a document and wishes to select the USPTO page, he uses a pointing device such as a mouse to point to the underlined phrase, causing the
25 terminal to extract the associated address and transmit it for selection of a new document.

The function of the apparatus 1 of Figure 1 is, in brief as follows:

- (a) to receive HTML documents from the server 4;
- (b) to synthesise an audio signal reciting the text contained in the
30 document and transmit it via the line 3 to the user at 2;
- (c) to recognise spoken replies from the user;
- (d) to recognise the replies from the user as indicating a selection of a further document;

(e) to transmit the address of that document to the server 4.

Figure 3 shows the apparatus 1 in more detail. It contains a network interface 10 which comprises a modem for connection to the link 5, and a processor programmed with software to transmit addresses via the modem to the server and receive documents from the server. This software differs from conventional browser software such as Netscape only in that (a) it receives addresses via a connection 11 rather than having them typed in at a keyboard or selected using a mouse and (b) it outputs the received text directly to a file or buffer 12 which can be accessed via a connection 13.

Suppose that a document has been received by the interface 10 and is stored in the buffer 12. A first portion of text is read out and a correspondingly coded signal is output on the line 13. The actual amount of the text output could rely on punctuation characters included in the text, for example up to the first (or second etc.) full stop, or up to the first paragraph mark.

This is received by a text pre-processing unit 14 which serves to delete unwanted control information, and forward it to a conventional text-to-speech synthesiser 15. This produces an audio signal corresponding to the portion of text, which is transmitted over the telephone line 3 to the user at 2.

The portion of text is also copied to a buffer 16. This is shown as coming from a second output of the pre-processing unit 14, since whilst the unit 14 removes from the text sent to the synthesiser 15 all format and control information (i.e. the characters < and > and anything within them), the text sent to the buffer 16 still includes the link address commands (e.g. but omits all other formatting and control information.

If desired, one could allow selected markings to pass to the synthesiser, for example so that bold type could be more heavily stressed, but this is entirely optional.

Although the link addresses are stored in the buffer 16, they are removed by further text processing 17 before forwarding the text to a recognition network generator 18 which is connected to a speech recogniser 19.

The recogniser 19 is connected to receive audio signals from the telephone line 3, so that responses from the user at 2 may be recognised. The

recogniser may have permanent programming to enable it to recognise some standard command words for control of the system; however its primary purpose is to match the user's response to the source text which has just been spoken by the synthesiser 15; more particularly to identify that part of the source text

- 5 present in the buffer 16 which most closely resembles the user's response.

Thus the function of the recognition network generator 18 is to derive, from the text input to it, parameters for the recogniser defining a vocabulary and grammar corresponding to this task.

In this example, it is assumed that the output of the recognser is a text
10 string corresponding to the matched portion of text (or command word). This output representing the user's response is taken to be a request for a further document information, and the next task is to identify this by locating the text string in the buffer 16 and returning the link address contained within in; or if there is none, returning the nearest link address stored in the buffer. This function (to
15 be discussed in more detail below) is performed by a link resolve unit 20 which outputs the link address to the interface 10, which transmits it to the database server 4 as a request for a further document. If however the link represents a position in the current document, then this is recognised and a command issued to the buffer 12 to read text from a specified point.

20 Control functions - for example if the user wishes to move on to the next (or preceding) paragraph of the document currently stored in the buffer 12, or to return to some default document, or to terminate the connection - could be performed using the telephone keypad, but preferably is achieved by designating certain words as control words (e.g. More, Back, Home, Quit) stored as a
25 permanent vocabulary in the recogniser 19 and received by a control unit 21 which, upon receiving one of these words along, then issues appropriate instructions to the buffer 12 and/or interface 10.

By way of further explanation of the operation of the apparatus, and in particular of the link resolver 20, consider a situation where the buffer 12 is loaded
30 with a document as shown in Figure 4A. The appearance of this document were it to be displayed on a visual display unit is as in Figure 4B.

Suppose that the buffer 12 is set up to output one paragraph at a time; suppose further that the user has already heard the title and asked for "More", the

buffer 12 outputs the next paragraph ("Welcome forests") to the text pre-processor as shown in Figure 4C.

Suppose now the user says "the Amazon basin". The recogniser 19 matches the speech signal and outputs the text string "Amazon basin", whereupon 5 the link resolver 20 searches in the buffer 16 for this text string, finds that it is attached to the link address <http://www/amazon.basin>, reads out this address and forwards it to the interface 10 which transmits it to the database server 4 to call up another page.

Naturally the user cannot know which expressions have link addresses 10 attached and to cater for the possibility of him/her uttering some other words, the link resolver operates according to the flowchart shown in Figure 5. In a first test 30, it is determined whether the matched source text is, or contains a link. "Amazon basin", "birds in the Amazon basin" or even "basin many of" would pass this test. In this case, the link address in question is chosen at 31. Otherwise a 15 second test 32 is performed to establish whether the matched source text lies in a sentence which contains a link; "one thousand species" for example would fall into this category. In this case the address in that sentence (or, if more than one, the one nearest to the matched source text) is chosen. Otherwise the nearest link to the matched source text is chosen, for example by counting the number of words 20 (or the number of characters) from the matched text to the next link above and below it in the buffer, and choosing the link with the lower count. A more complex algorithm could examine the nearest links above and below the matched text for the degree of semantic similarity to the matched text and choose the more similar.

25 In a refinement, one could weight this choice to take account of punctuation, for example by increasing by (e.g.) 10 words the count when crossing a paragraph boundary.

The HTML language also permits links to other parts of the current document - as shown in Figure 4A for the British Wildlife Society. Upon 30 recognition of this name by the recogniser, the address "#3224" would be recognised by the link resolver as an internal address and forwarded not to the interface 10 but to the buffer 12 to cause readout of a paragraph from a point in the document specified by the address.

The operation of the recognition network generator 18 may now be discussed further. There are essentially two components to the setting up of a recogniser for a given function. First, defining its vocabulary, and second, defining its grammar. The vocabulary is a question of ensuring that the recogniser has a

5 set of models or templates, typically one for each of the words to be recognised - that is, one for each of the words (other than link addresses) present in the buffer

16. Vocabulary generation for this purpose may use any of the conventional methods. Typically this is done by using a recogniser preprogrammed with a set of sub-word models (e.g. one per phoneme) and processing each word delivered from

10 the buffer, in similar manner to the operation of a text-to-speech synthesiser, to generate a word template by concatenation of the appropriate sub-word models. Alternatively the recogniser may have a standard store of word models which can be retrieved when the corresponding words are received from the buffer 16, though to accommodate proper names and other words not in the standard set the

15 sub-word concatenation method would usually be employed as well.

The grammar of a recogniser is a set of stored parameters which define what word sequences are permissible; for example, considering the buffer contents shown in Figure 4A whilst "Amazon basin" is a word sequence which is useful to recognise "basin Amazon" is not. One possibility is to allow (as sequences for

20 matching against the user's utterance) any number of words from 1 upwards, but only in the sequence in which they appear in the buffer. Figure 6 shows this represented graphically (for a portion only of the text) where 40 represents a start node of a recognition "tree", 41 represents an end node, 42 represents word models and the lines 43 represent allowable paths so.

25 It would be possible to include a network of 'carrier phrases' as shown in Figure 7 so that the user could say sentences such as "Tell me more about the Amazon Basin please". Alternatively a garbage or sink model (Fig. 8) could be included at the beginning and end of the network to allow any speech to surround the echoed phrase.

30 In another embodiment the recogniser could simply allow any of the words on the page to be uttered in any order as shown in Figure 9. The accuracy of such a recogniser would not be as high as those shown in Figures 6 to 8, but if

statistical constraints based on the contents of the HTML page were incorporated in the recognition process a working system could be created.

Returning briefly to Figure 3, in this embodiment it has been assumed that the recogniser returns, as a "label" representing its recognition result, the relevant 5 part of the actual text string supplied to the recognition network generator 18 by the buffer 16, and the link resolver 20 matches this string against the buffer contents to locate the desired links. Whilst this may be convenient to permit use of a conventional unit for the recogniser 16, a way of speeding up the operation of the link resolver would be to set up the recogniser to return some parameter 10 enabling faster access to the buffer, for example pointer values giving the addresses in the buffer 16 of the first and last characters of the matched source text string.

Although only one server is shown in Figure 1, of course there could be others, and the transmitted link address could well be destined for a different 15 server from the one sending the document from which it was obtained.

This embodiment presupposes that the source text carries hyperlink addresses; however it is also possible to operate this system without embedded addressed of this kind. For example one could transmit to the database server 20 coordinates to identify the point in a (or range of) the source text at which the match occurred. In the case of connectionless service such as the Internet, it would be necessary to concatenate this information with the address of the server before transmitting it.

It was mentioned earlier that the text preprocessor 14 could be arranged to pass certain markings through to the synthesiser 15 to allow bold type to be 25 emphasised. Similarly, it would be possible for the preprocessor to pass the hyperlink markings <a>... (albeit without the addresses) and arrange the synthesiser to respond to these by applying an emphasis, or even switching to a different voice (for example a male instead of female voice) from that used for the remainder of the text. With this expedient, in an alternative embodiment, one can 30 simplify the speech recogniser vocabulary to include only the link words, though it is still preferred to operate the recogniser as described above, against the possibility that the user may not always accurately recollect which words were spoken with the emphasis (or different voice).

CLAIMS

1. An interface for a voice interactive service comprising:
a speech synthesiser to receive coded signals representing sequences of
5 words and to generate audio signals corresponding thereto for output;
speech recognition means connected to receive the said coded signals
and operable upon receipt of a speech signal to be recognised to identify that part
of the word sequence represented by the coded signals which most resembles the
speech signal to be recognised.
10
2. An interface according to Claim 1 in which the coded signals include link
signals identifying one or more words of a sequence which represent links to
further information, and the apparatus is operable to select from the coded signals
a link signal which is in or adjacent to the identified resembling part of the
15 sequence.
3. An interface according to Claim 2 including a communications interface
connected to receive the coded signals from a remote source and to transmit the
selected link signal to the same or another remote source for requesting further
20 coded signals.
4. An interface according to Claim 2 or 3 including a buffer for storing the
coded signals, wherein:
 - (a) the interface is so operable that the speech synthesiser can generate
25 audio signals corresponding to a portion only of the coded signals stored in the
buffer and the recogniser thereupon identifies that part of the word sequence
which is represented by said portion of the coded signals which part most
- resembles the speech signal to be recognised;
and
 - (b) the interface includes control means responsive to a link signal
30 which identifies a further portion of the coded signals stored in the buffer to
transmit that further portion to the synthesiser and recognition means.

5. An interface according to any one of the preceding claims including a telephone line interface whereby the generated audio signals and received speech signals may respectively be sent to and received from a remote user.

5 6. An interface for a voice interactive service as herein described with reference to the accompanying drawings.

7. A method of operating a voice interactive service comprising

(a) receiving coded signals representing a sequence of words and

10 synthesising audio signals corresponding thereto for output;

(b) receiving a speech signal and identifying by means of a speech recogniser that part of the word sequence represented by the coded signals which most resembles the received speech signal; and

(c) using the recognition result to select a further sequence of words.

15

8. A method according to Claim 7 in which the coded signals include link signals identifying one or more words of a sequence which represent links to further information, and step (c) includes selecting from the coded signals a link signal which is in or adjacent the identified resembling part of the sequence.

20

9. An interface for a voice interactive service comprising:

a speech synthesiser to receive coded signals representing sequences of words and to generate audio signals corresponding thereto for output, in which the coded signals include link signals identifying one or more words of a sequence

25 which represent links to further information, the synthesiser being responsive to receipt of the link signals to utter the words so identified in a different manner from words not so identified; and

speech recognition means connected to receive at least those of the coded signals which represent link-representing words and operable upon receipt of
30 a speech signal to be recognised to identify which of the link-representing words most resemble the speech signal to be recognised.

1/9

Fig.1.

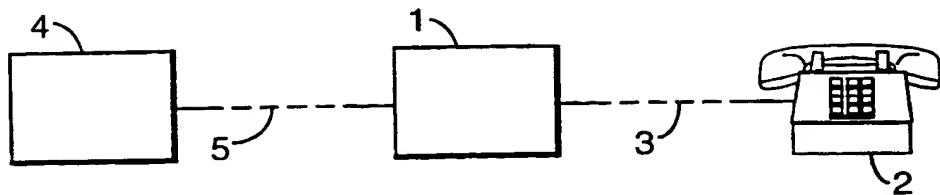


Fig.2A.

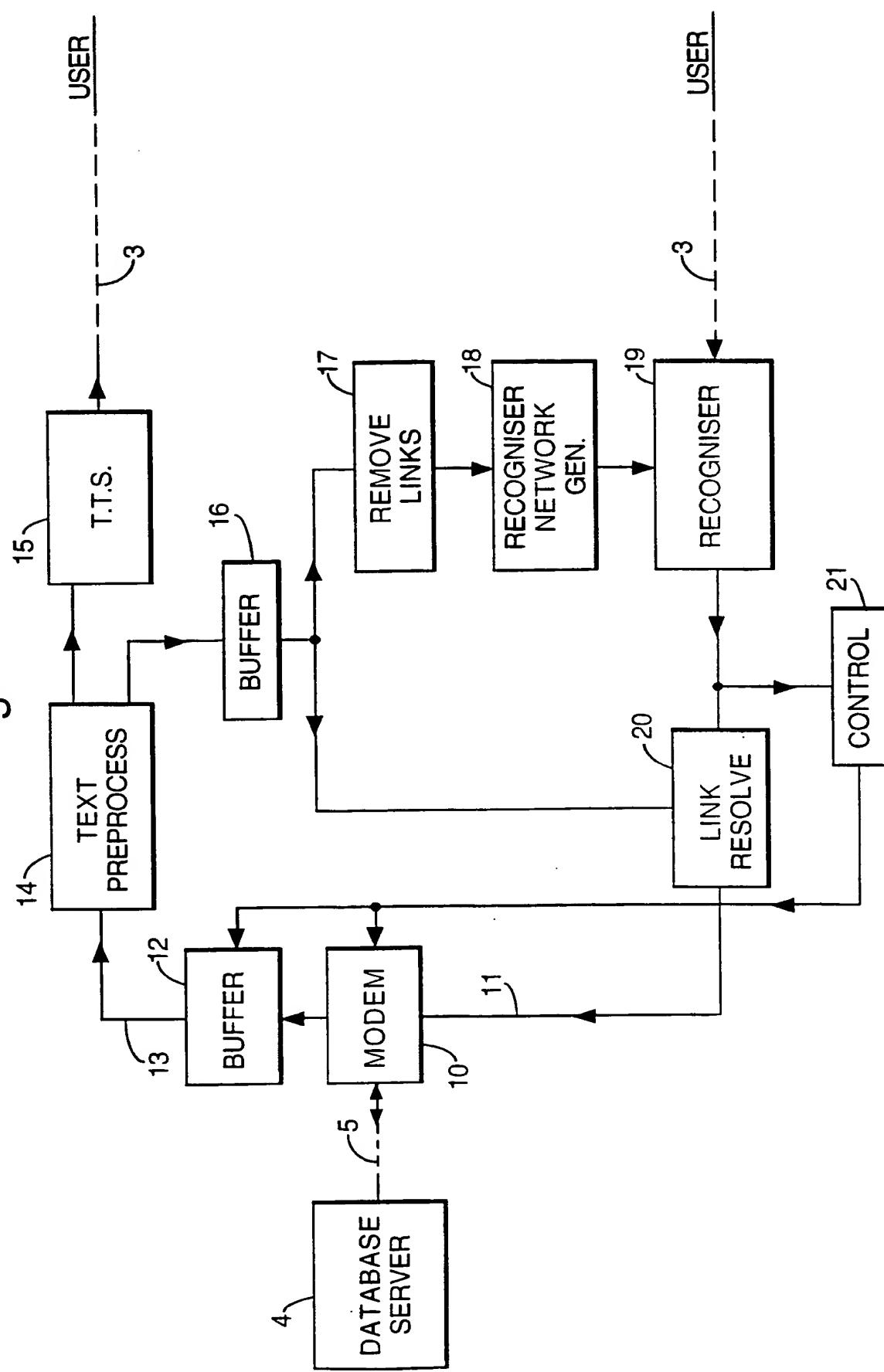
```
<li><b>Patent Office Sites</b>
<ul>
<li><a href="http://www.uspto.gov">US Patent and Trademark Office</a>
<li><a href="http://www.epo.co.at/epo">European Patent Office</a>
<li><a href="http://www.uspto.gov/wipo.html">World Intellectual Property
Organisation</a>
<li><a href="http://www.jpo-miti.go.jp">Japanese Patent Office</a>
<li><a href="http://www.netwales.co.uk/ptoffice/index.htm">UK Patent Office</a>
</ul>
```

Fig.2B.

PATENT OFFICE SITESUS Patent and Trademark OfficeEuropean Patent OfficeWorld Intellectual Property OrganisationJapanese Patent OfficeUK Patent Office

2/q

Fig.3.



3/9

Fig.4A.

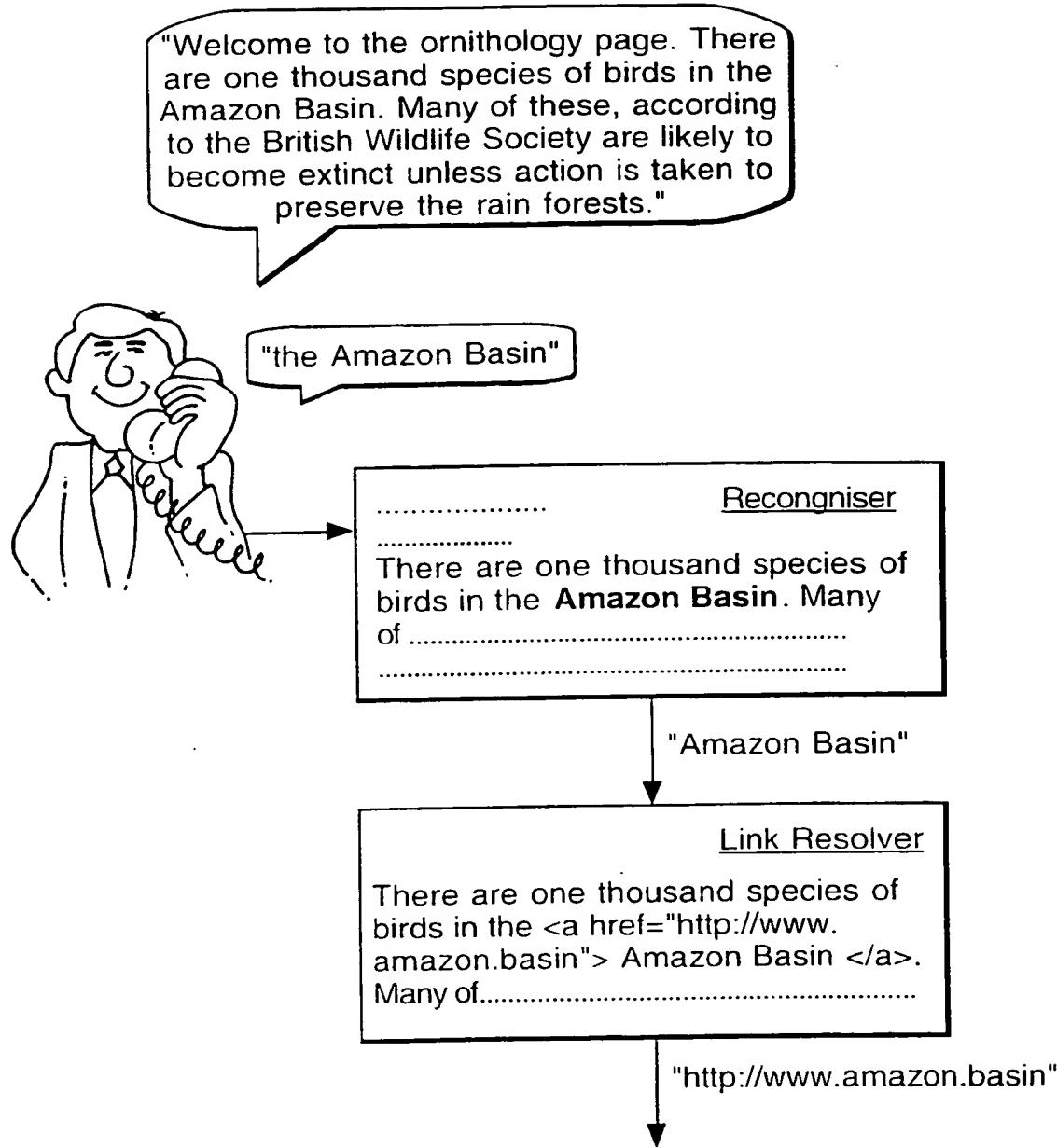
 Ornithology <p>Welcome to the ornithology page. There are over one thousand species of birds in the Amazon basin. Many of these, according to the British Wildlife Society are likely to become extinct unless action is taken to preserve their environment, the rain forests. <p> Other areas in which birds.....

Fig.4B.

Ornithology
Welcome to the **ornithology page**. There are over one thousand species of birds in the Amazon basin. Many of these, according to the British Wildlife Society are likely to become extinct unless action is taken to preserve their environment, the rain forests.
Other areas in which birds.....

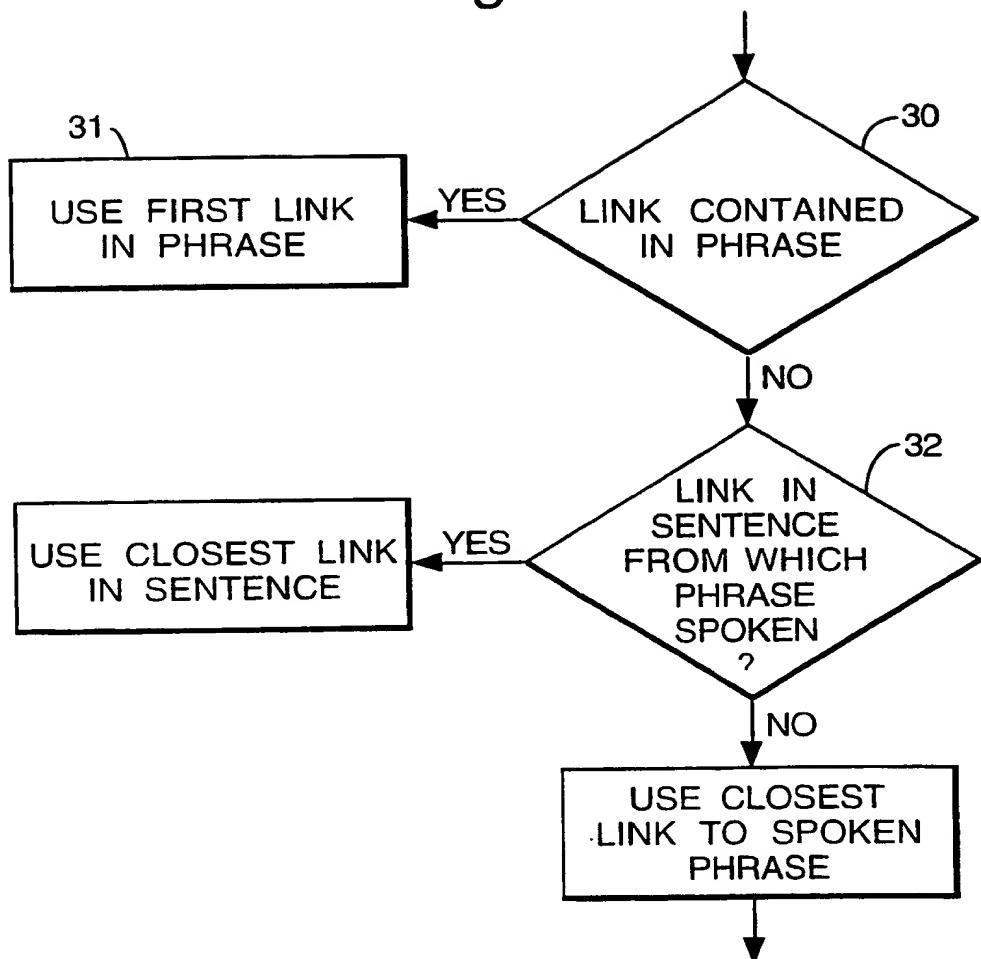
4/q

Fig.4C.



5/9

Fig.5.



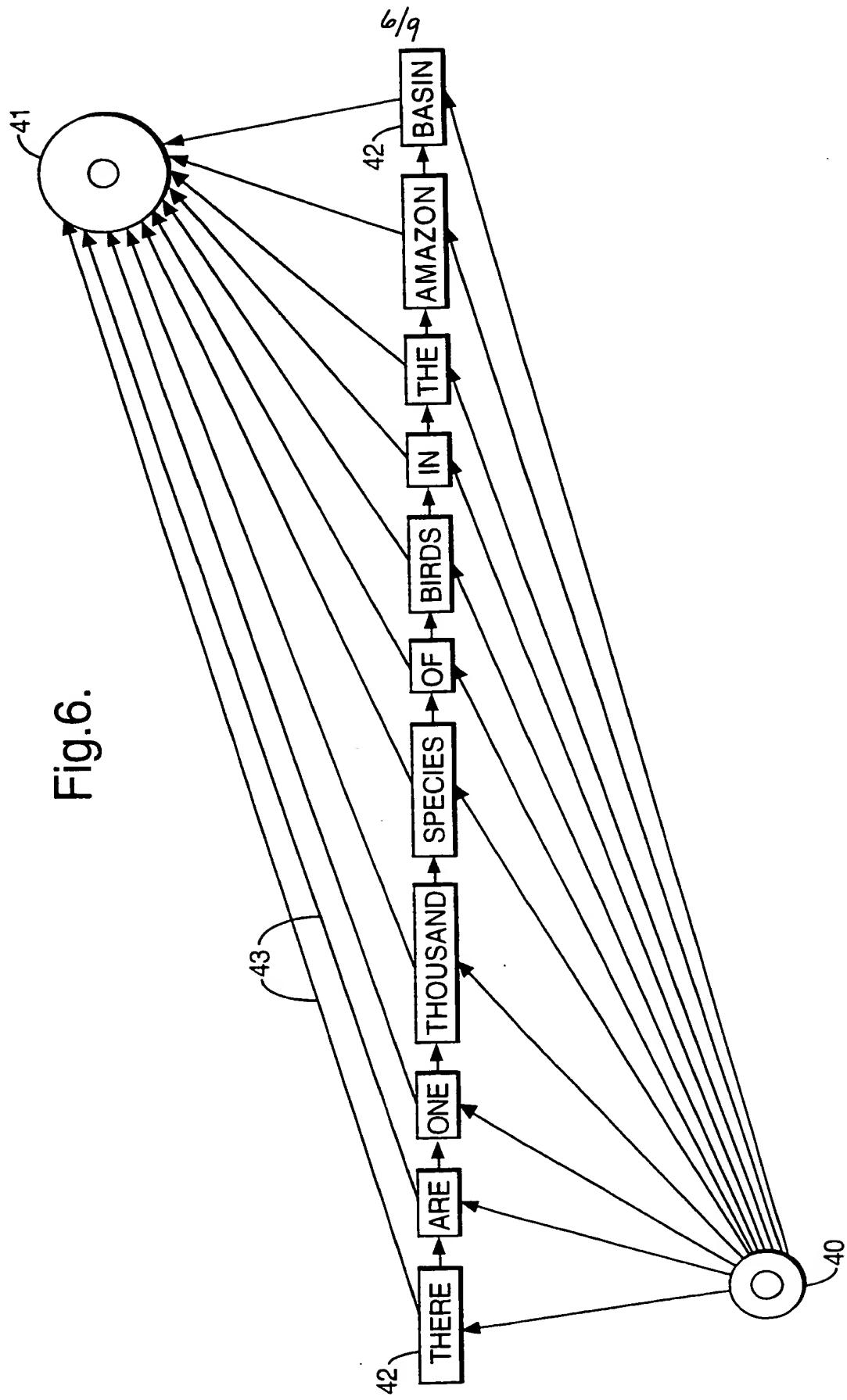
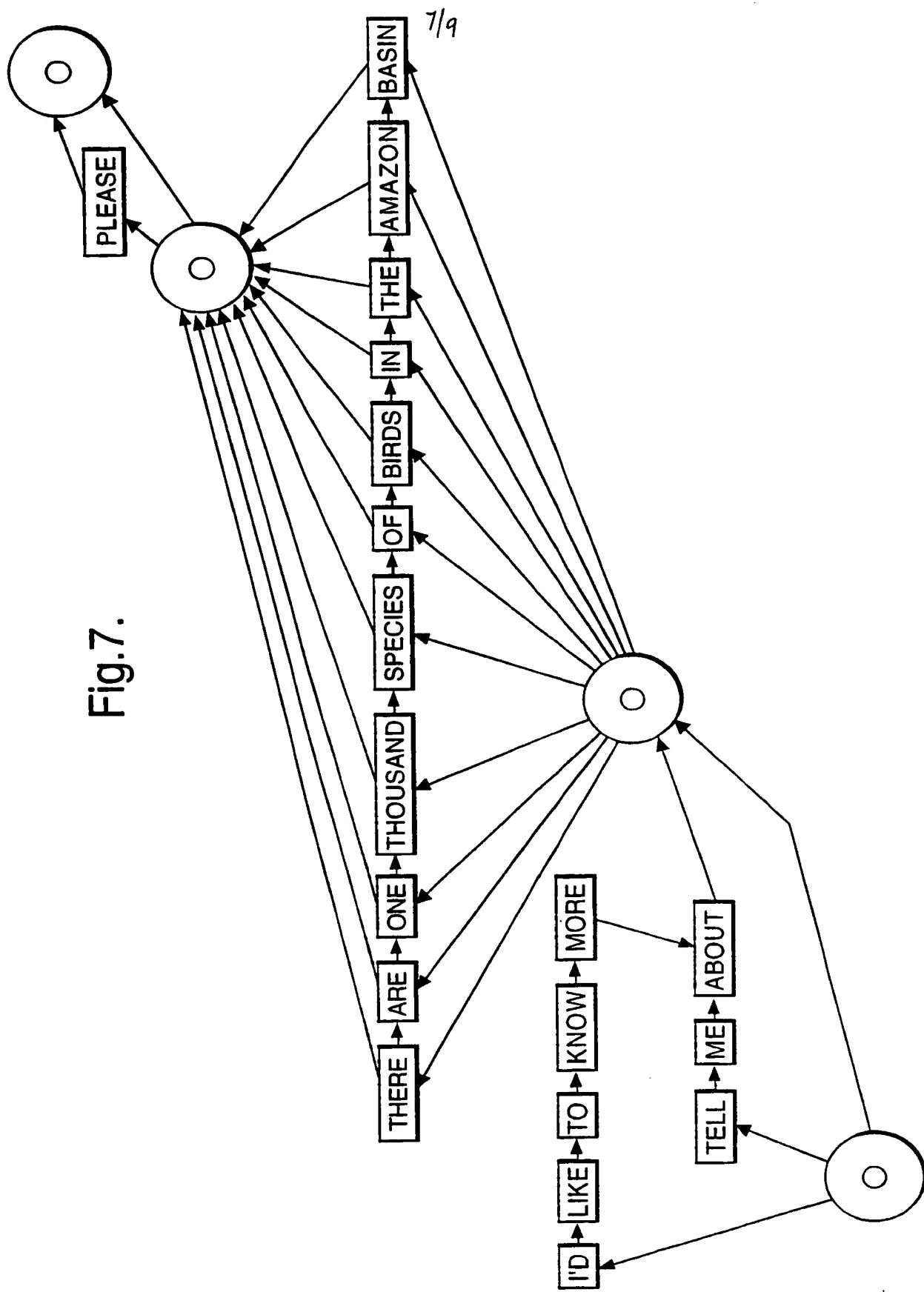
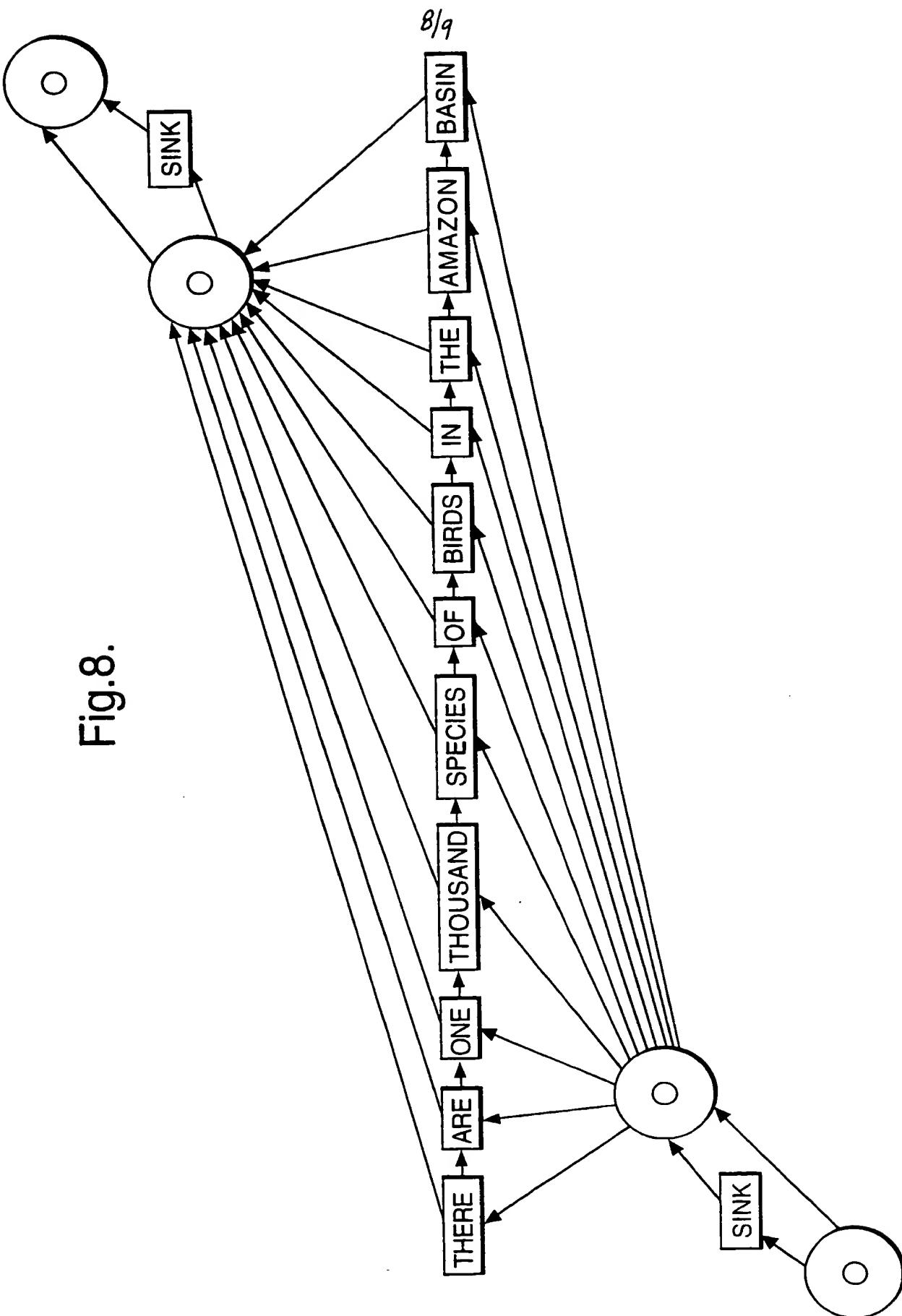


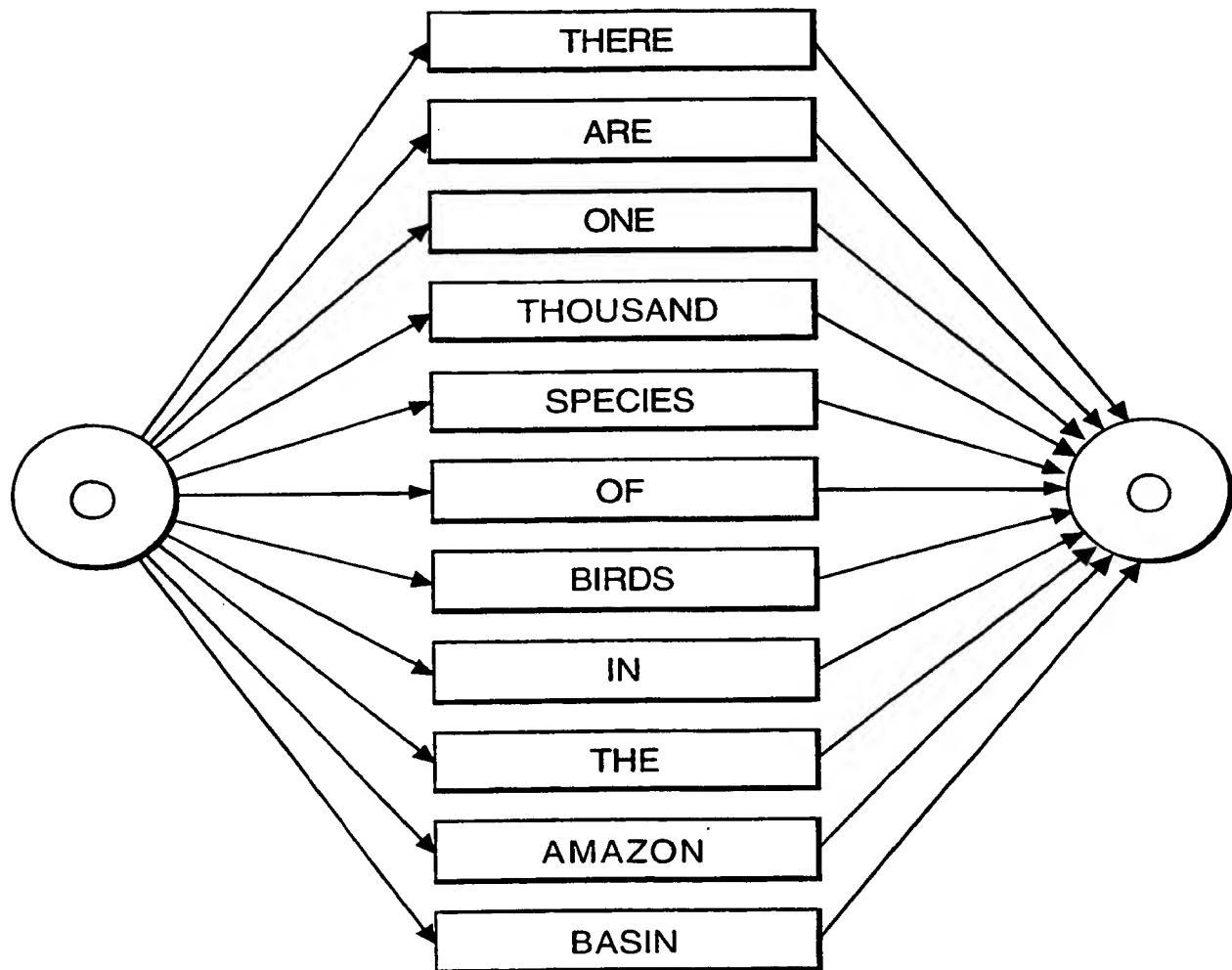
Fig. 7.





9/9

Fig.9.



INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 98/00194

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 H04M3/50 H04M7/00 G06F17/30

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 H04M G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ATKINS D L , BALL T , BARAN T R , BENEDICT M A , COX K C , LADD D A , MATAGA P A , PUCHOL C , RAMMING J C , REHOR K G , TUCKEY C.: "INTEGRATED WEB AND TELEPHONE SERVICE CREATION" BELL LABS TECHNICAL JOURNAL, vol. 2, no. 1 January 1997, USA, pages 19-35, XP002036350 see the whole document ---	1-9
X	EP 0 735 736 A (AT & T IPM CORP) 2 October 1996 see the whole document ---	1,5,7 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

" Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

8 June 1998

Date of mailing of the international search report

16/06/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Megalou, M

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 98/00194

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RICCIO A ET AL: "VOICE BASED REMOTE DATA BASE ACCESS" PROCEEDINGS OF THE EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY (EUROSPEECH), PARIS, SEPT. 26 - 28, 1989, vol. 1, 26 September 1989, TUBACH J P; MARIANI J J, pages 561-564, XP000209922 see the whole document	1,5,7
A		4
X	RABINER L R: "The impact of voice processing on modern telecommunications" SPEECH COMMUNICATION, vol. 17, no. 3-4, November 1995, pages 217-226, XP000641894 see the whole document	1,5,7
A	US 5 572 625 A (RAMAN T V ET AL) 5 November 1996 see abstract see figure 1 see column 32, line 35 - column 36, line 59	1-9
P,X	GB 2 307 619 A (POLLITT ALEXANDER JAMES) 28 May 1997 see the whole document	1-3,5, 7-9
P,X	WO 97 23973 A (UNIV RUTGERS ;IMIELINSKI TOMASZ (US); VIRMANI AASHU (US)) 3 July 1997 see page 10, line 16 - page 19, line 16	1-9
P,X	WO 97 40611 A (AT & T CORP) 30 October 1997 see abstract see page 8, line 3 - page 10, line 3 see figures 2-6	1-9
A	PAGE J H ET AL: "THE LAUREATE TEXT-TO-SPEECH SYSTEM - ARCHITECTURE AND APPLICATIONS" BT TECHNOLOGY JOURNAL, vol. 14, no. 1, 1 January 1996, pages 57-67, XP000554639 see the whole document	1-9
A	TAKAHASHI J ET AL: "INTERACTIVE VOICE TECHNOLOGY DEVELOPMENT FOR TELECOMMUNICATIONS APPLICATIONS" SPEECH COMMUNICATION, vol. 17, no. 3/04, November 1995, pages 287-301, XP000641897 see the whole document	1,7

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 98/00194

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP 0735736 A	02-10-1996	US JP	5724481 A 8320696 A	03-03-1998 03-12-1996
US 5572625 A	05-11-1996		NONE	
GB 2307619 A	28-05-1997		NONE	
WO 9723973 A	03-07-1997	AU	1566497 A	17-07-1997
WO 9740611 A	30-10-1997	EP	0834229 A	08-04-1998

THIS PAGE BLANK (USPTO)